2008/34

■

Polynomial-time computation of the joint spectral radius for some sets of nonnegative matrices

Vincent D. Blondel and Yurii Nesterov

CORE
Voie du Roman Pays 34
B-1348 Louvain-la-Neuve, Belgium.
Tel (32 10) 47 43 04
Fax (32 10) 47 43 01
E-mail: corestat-library@uclouvain.be
http://www.uclouvain.be/en-44508.html

CORE DISCUSSION PAPER
2008/34

Polynomial-time computation of the joint spectral radius for some sets of nonnegative matrices

Vincent D. BLONDEL [1] and Yurii NESTEROV[2]

May 2008

Abstract

We propose two simple upper bounds for the joint spectral radius of sets of nonnegative matrices. These bounds, the *joint column radius* and the *joint row radius*, can be computed in polynomial time as solutions of convex optimization problems. We show that for general matrices these bounds are within a factor $1/n$ of the exact value, where n is the size of the matrices. Moreover, for sets of matrices with independent column uncertainties of with independent row uncertainties, the corresponding bounds coincide with the joint spectral radius. In these cases, the joint spectral radius is also given by the largest spectral radius of the matrices in the set. As a byproduct of these results, we propose a polynomial-time technique for solving Boolean optimization problems related to the spectral radius. We also consider economics and engineering applications of our results which were never considered practice due to their intrinsic computational complexity.

Keywords: joint spectral radius, joint column radius, joint row radius, nonnegative matrices, asynchronous systems, convex optimization, Leontief model.

---

[1] Division of Applied Mathematics, Université catholique de Louvain, Belgium. E-mail: Vincent.blondel@uclouvain.be
[2] CORE and INMA, Université catholique de Louvain. E-mail: yurii.nesterov@uclouvain.be. This author is also member of ECORE, the newly created association between CORE and ECARES.

The research results presented in this paper have been supported by a grant "Action de recherche concertée ARC 04/09-315" from the "Direction de la recherche scientifique – Communauté française de Belgique".

This paper presents research results of the Belgian Program on Interuniversity Poles of Attraction initiated by the Belgian State, Prime Minister's Office, Science Policy Programming. The scientific responsibility is assumed by the authors.

1 Introduction

A discrete-time switching linear system generates trajectories of points

$$x_{k+1} = A_k x_k, \quad x_0 \in R^n, \tag{1.1}$$

with the matrices A_k taken in some *uncertainty set* $\mathcal{M} \subset R^{n\times n}$. The worst-case growth rate of these trajectories can be characterized by a joint spectral radius. The joint spectral radius (JSR) of the set of matrices $\mathcal{M}$ is the smallest value $\rho \geq 0$ such that for every trajectory there is some constant C for which

$$\|x_k\| \leq C\,\rho^k$$

for all k. This optimal ρ provides valuable information about the switching linear system. In particular, the trajectories of the switching system all converge to the origin if and only if $\rho < 1$. In [18] the joint spectral radius of $\mathcal{M}$ is defined in the following equivalent form:

$$\rho(\mathcal{M}) = \limsup_{k\to\infty} \max\{\|A\|^{1/k} : A \text{ is a product of length } k \text{ of matrices in } \mathcal{M}\}. \tag{1.2}$$

During the last decade, the joint spectral radius has proved useful in a number of application contexts, including wavelets [7,9], capacity of codes [5,14], sensor networks [8,13], combinatorics on words [12], autoregressive models, Markov chains, etc. Unfortunately, the joint spectral radius of a set of matrices is notoriously difficult to compute and to approximate. In fact, even for the case of two matrices of dimension 47×47 that have non-negative rational entries, the problem of checking the inequality $\rho \leq 1$ is algorithmically undecidable (see [4,6]), and it is still unknown if the problem $\rho < 1$ is algorithmically decidable. Today, the list of matrix sets $\mathcal{M}$ that have polynomially computable JSR is desperately small: the list includes the case where $\mathcal{M}$ contains only symmetric matrices, only triangular matrices of identical orientation, or two symmetric matrices, $\mathcal{M} = \{A, A^T\}$. In this last case the joint spectral radius is given by the largest singular value of the matrix.

In recent years most research efforts have concentrated on finding reasonable approximations for the JSR (see, e.g., [2,3,16,17]). But all these approximation suffer from an intrinsic limitation since it is known that the problem of computing the JSR of two matrices with binaries entries is NP-hard and that, unless P=NP, there is no approximation algorithm that is polynomial with respect to the accuracy [19]. In all known approximation schemes, even for sets with only two matrices, the complexity of the computations grows exponentially with respect to the required accuracy.

In this paper, we propose two upper bounds for the JSR of arbitrary sets of nonnegative matrices, which are both within a factor $1/n$ of the exact value. These bounds, the *Joint Column Radius* (JCR) and the *Joint Row Radius* (JRR), can be computed in polynomial time as solutions to convex optimization problems.

We also consider the special case for which the set of matrices $\mathcal{M}$ has independent column, or row, uncertainties. In terms of the switching systems (1.1) the row independent uncertainty situation corresponds to systems for which at every iteration all entries of the state vector x_k are updated, and the i-th entry is updated by choosing one of the elements in $\{q^T x_t : q \in Q_i\}$.

In particular, this includes the situation of asynchronous linear systems for which at every iteration only some of the state entries are updated according to a linear transformation and the others are kept unchanged.

It appears that in this special case, the JRR coincides with the JSR and with the largest spectral radius of the matrices in the set. As a byproduct of this result, we get a possibility to solve in polynomial time some boolean optimization problems related to the spectral radius. Another interesting consequence of our results is the quasi-convexity of the spectral radius of a matrix with nonnegative matrices in each column, when all other columns are fixed.

This paper is organized as follows. In the next section, we define the joint column (row) radius. We show that this value is the solution of a convex optimization problem and we establish an upper and a lower bound for its quality as an approximation to the JSR. Then, in Section 3, we show that for sets of matrices that have independent column (or row) uncertainty sets, the JCR (JRR) coincides with the JSR. Moreover, it appears that in the uncertainty set there is always a critical matrix whose spectral radius coincides with the JSR. This surprising property is valid even for discrete uncertainty sets.

In the last section, we discuss several applications of our results.

Notation. The entries of a (column) vector $x \in R^n$ are denoted by $x = (x^{(1)}, \ldots, x^{(n)})^T$. For x and y in R^n we denote by $\langle x, y\rangle$ their standard inner product:

$$\langle x, y \rangle \;=\; \sum_{i=1}^{n} x^{(i)} y^{(i)} \;=\; x^T y.$$

The set of square matrices is denoted by M_n and the set of square matrices with nonnegative entries is denoted by M_n^+. For a vector $x \in R^n$, we denote by $D(x) \in M_n$ the diagonal matrix with the vector x as its diagonal, and by $e^x \in R^n$ the vector with coordinates $e^{x^{(i)}}$. Further, $\mathbf{1}$ denotes the vector of all ones, $\mathbf{0}$ denotes the vector of all zeros, and e_i denotes the ith coordinate vector in R^n. Finally, Δ_n denotes the standard simplex in R^n, and $|\mathcal{M}|$ the cardinality of the set $\mathcal{M}$.

2 Joint column (row) radius

Let $\rho(A)$ be spectral radius of the matrix A, i.e., the largest magnitude of its eigenvalues. According to the Perron-Frobenius Theorem, an irreducible[1] nonnegative matrix $A \in M_n^+$ has a unique eigenvector u (up to scalar multiplication) such that

$$A^T u \;=\; \rho(A)\, u,$$

and all components of the vector u are then positive. Let the column decomposition of A be given by $A = (a_1, \ldots, a_n)$. It is well known (see, e.g., [10]) that the spectral radius of a

[1] A matrix A is reducible if there is a permutation matrix P for which:

$$PAP^T \;=\; \begin{pmatrix} F & 0 \\ G & H \end{pmatrix}. \qquad (2.1)$$

The matrix is irreducible if no such permutation exists.

nonnegative matrix admits the following representation:

$$\rho(A) \;=\; \inf_{u>0} \max_{1\le i\le n} \tfrac{\langle a_i, u\rangle}{u^{(i)}}. \tag{2.2}$$

Changing the variables $u = e^x$, we obtain

$$\rho(A) \;=\; \inf_{x\in R^n} \left[\phi_A(x) \stackrel{\mathrm{def}}{=} \max_{1\le i\le n} \langle a_i, e^x\rangle \cdot e^{-x^{(i)}} \right]. \tag{2.3}$$

Note that the objective function $\phi_A(x)$ in this problem is convex. If A is irreducible, then the problem (2.3) has a unique solution ray spanned by direction $\mathbf{1} \in R^n$. We denote by $x(A)$ the point of the optimal ray satisfying the equation

$$\langle \mathbf{1}, x(A)\rangle \;=\; 0,$$

and $u(A) \stackrel{\mathrm{def}}{=} e^{x(A)} > \mathbf{0}$. Note that $A^T u(A) = \rho(A)\cdot u(A)$.

Representation (2.2) explains the role of the spectral radius in estimating the rate of growth/decrease of the powers of positive matrices. Indeed, for an arbitrary point $x_0 \in R^n_+$ define the sequence

$$x_k \;=\; A^k x_0, \quad k \ge 1.$$

Then

$$\begin{aligned} \langle u(A), x_{k+1}\rangle \;&=\; \langle u(A), Ax_k\rangle \;=\; \langle A^T u(A), x_k\rangle \\ &=\; \langle D^{-1}(u(A)) A^T u(A), D(u(A)) x_k\rangle \\ &\stackrel{(2.2)}{\le}\; \langle \rho(A)\cdot \mathbf{1}, D(u(A)) x_k\rangle \;=\; \rho(A)\cdot\langle u(A), x_k\rangle. \end{aligned} \tag{2.4}$$

It is interesting that exactly the same reasoning can be used in the analysis of switching systems.

Consider a compact set $\mathcal{M}$ of nonnegative matrices. We define the *joint column radius* (JCR) of $\mathcal{M}$ as follows:

$$\sigma(\mathcal{M}) \;=\; \inf_{x\in R^n} \max_{A\in\mathcal{M}} \phi_A(x). \tag{2.5}$$

In this expression, all functions $\phi_A(x)$ are convex in x and so the function $\max_{A\in\mathcal{M}} \phi_A(x)$ is also convex. Since the JCR is a solution of the convex minimization problem (2.5), it can be computed in polynomial time by standard convex optimization algorithms.

We can provide JCR with another interesting interpretation. Denote $\widehat{\mathcal{M}} = \mathrm{Conv}\,(\mathcal{M})$ and consider the following optimization problem

$$\max_{A\in\widehat{\mathcal{M}}} \rho(A). \tag{2.6}$$

The inequality

$$\max_{A\in\widehat{\mathcal{M}}} \rho(A) \;\le\; \rho(\mathcal{M}). \tag{2.7}$$

was proved as Proposition 15 in [3] for a finite collection $\mathcal{M}$ and by Carateodory Theorem it can easily be extended to arbitrary convex uncertainty sets.

Note that

$$\max_{A \in \widehat{\mathcal{M}}} \rho(A) \stackrel{(2.3)}{=} \max_{A \in \widehat{\mathcal{M}}} \inf_{x \in R^n} \phi_A(x) \le \inf_{x \in R^n} \max_{A \in \widehat{\mathcal{M}}} \phi_A(x)$$

$$= \inf_{x \in R^n} \max_{A \in \mathcal{M}} \phi_A(x) = \sigma(\mathcal{M}).$$

Thus, we can treat $\sigma(\mathcal{M})$ as a value of the usual *Lagrangian relaxation* of the nonconvex optimization problem (2.6). Note that

$$\max_{A \in \mathcal{M}} \rho(A) \le \max_{A \in \widehat{\mathcal{M}}} \rho(A) \le \sigma(\mathcal{M}).$$

In Section 3 we will discuss nontrivial situations when these inequalities can be replaced by equalities.

A quantity similar to the JCR can be introduced for the set of transposed matrices

$$\mathcal{M}^T \stackrel{\rm def}{=} \{A^T, \ A \in \mathcal{M}\}.$$

Namely, we can define the *joint row radius* (JRR) of the set $\mathcal{M}$ as follows:

$$\sigma_T(\mathcal{M}) = \sigma(\mathcal{M}^T). \tag{2.8}$$

Since $\rho(\mathcal{M}) = \rho(\mathcal{M}^T)$, the above discussion applies to the JRR as well. However, in general we have $\sigma_T(\mathcal{M}) \neq \sigma(\mathcal{M})$. In the remaining part of the paper we will work mainly with the JCR, without mentioning that all our results can also be applied to the JRR.

Let us now prove that the JCR provides good upper and lower bounds for the JSR. First of all, we need to prove two technical lemmas.

Lemma 1 *Consider the following perturbation of the uncertainty set* $\mathcal{M}$*:*

$$\mathcal{M}_\epsilon \stackrel{\rm def}{=} \{A + \epsilon \mathbf{1}\mathbf{1}^T, \ A \in \mathcal{M}\}, \quad \epsilon \ge 0.$$

Then $\lim\limits_{\epsilon \downarrow 0} \sigma(\mathcal{M}_\epsilon) = \sigma(\mathcal{M})$*.*

Proof:
Indeed,

$$\begin{aligned} \sigma(\mathcal{M}) \le \sigma(\mathcal{M}_\epsilon) &= \inf_{u > \mathbf{0}} \max_{A \in \mathcal{M}} \max_{1 \le i \le n} \left[\tfrac{1}{u^{(i)}} \cdot \langle a_i + \epsilon \mathbf{1}, u \rangle \right] \\ &\le \inf_{\substack{\langle \mathbf{1}, u \rangle = 1, \\ u > \mathbf{0}}} \max_{A \in \mathcal{M}} \max_{1 \le i \le n} \left[\tfrac{1}{u^{(i)}} \cdot \langle a_i, u \rangle + \tfrac{\epsilon}{u^{(i)}} \right] \\ &\le \inf_{\substack{\langle \mathbf{1}, u \rangle = 1, \\ u > \mathbf{0}}} \left[\xi_{\mathcal{M}}(u) + \epsilon F(u) \right] \stackrel{\rm def}{=} \tau(\epsilon), \end{aligned}$$

where $\xi_{\mathcal{M}}(u) = \max\limits_{A\in\mathcal{M}} \max\limits_{1\le i\le n} \left[\frac{1}{u^{(i)}}\cdot\langle a_i, u\rangle\right]$, and $F(u) = \max\limits_{1\le i\le n} \frac{1}{u^{(i)}}$ is a penalty function for the positive orthant. Since $F(u)$ is below bounded on its feasible set, by Theorem 1.3.2 in [15] we have[2]

$$\lim_{\epsilon\downarrow 0} \tau(\epsilon) \;=\; \inf_{\substack{\langle \mathbf{1},u\rangle=1,\\ u>\mathbf{0}}} \xi_{\mathcal{M}}(u) \;=\; \sigma(\mathcal{M}).$$

□

Lemma 2 *Let the elements of all matrices in $\mathcal{M}$ be positive. Then there exists a matrix $A_* = (A_1e_1,\ldots,A_ne_n)$ formed by some matrices $A_i \in \widehat{\mathcal{M}}$, $i=1,\ldots,n$, such that*

$$\rho(A_*) \;=\; \sigma(\mathcal{M}). \tag{2.9}$$

Proof:
Definition (2.5) of JCR can be written in the following form:

$$\begin{aligned}\sigma(\mathcal{M}) &= \inf_{x\in R^n} \max_{A\in\mathcal{M}} \max_{1\le i\le n} \left[e^{-x^{(i)}}\langle Ae_i, e^x\rangle\right] \\ &= \inf_{x\in R^n} \max_{1\le i\le n} \left[\psi_i(x) \stackrel{\mathrm{def}}{=} \max_{A\in\mathcal{M}} \left(e^{-x^{(i)}}\langle Ae_i, e^x\rangle\right)\right].\end{aligned} \tag{2.10}$$

Under conditions of the lemma, the infimum of the upper level minimization problem is attained at some point $x^* \in R^n$. Therefore, there exists a vector of optimal dual multipliers $\lambda_* \in \Delta_n$ such that

$$0 \;=\; \sum_{i=1}^n \lambda_*^{(i)} g_i^*, \quad g_i \in \partial\psi_i(x_*),$$

$$\partial\psi_i(x_*) \;=\; \left\{g = \nabla\left(e^{-x^{(i)}}\langle Ae_i, e^x\rangle\right),\ A_i \in \widehat{\mathcal{M}} :\ \langle Ae_i, e^{x_*}\rangle = e^{x_*^{(i)}}\psi_i(x_*)\right\}.$$

Moreover, there exist matrices $A_i^* \in \widehat{\mathcal{M}}$ such that

$$x^* \;=\; \arg\min_{x\in R^n} \max_{1\le i\le n} \left[e^{-x^{(i)}}\langle A_i^* e_i, e^x\rangle\right].$$

Hence, in view of (2.3), e^{x_*} is an eigenvector of the matrix $A_* = (A_1^*e_1,\ldots,A_n^*e_n)$, and $\rho(A_*) = \sigma(\mathcal{M})$. □

Now we can prove the main result of this section.

Theorem 1 *Let $\mathcal{M}$ be a compact set of nonnegative matrices. Then*

$$\tfrac{1}{p}\cdot\sigma(\mathcal{M}) \;\le\; \rho(\mathcal{M}) \;\le\; \sigma(\mathcal{M}), \tag{2.11}$$

where $p = \min\{n, |\mathcal{M}|\}$.

[2] In this theorem, the optimization problems are written in a min-form. However, all arguments work for the inf-form also.

Proof:
Let us fix an arbitrary $\epsilon > 0$. Then, by Lemma 2 there exist a matrix

$$A_*^\epsilon \;=\; (A_1^\epsilon e_1, \dots, A_n^\epsilon e_n), \quad A_i^\epsilon \in \widehat{\mathcal{M}}_\epsilon, \; i = 1, \dots, n,$$

which ensures equality $\rho(A_*^\epsilon) = \sigma(\mathcal{M}_\epsilon)$. Denote by m the number of different matrices in representation of A_*^ϵ:

$$\{A_i^\epsilon\}_{i=1}^n \;=\; \{B_j^\epsilon\}_{j=1}^m.$$

Then, $\frac{1}{m} A_*^\epsilon \leq B_\epsilon \stackrel{\mathrm{def}}{=} \frac{1}{m} \sum\limits_{j=1}^m B_j^\epsilon \in \widehat{\mathcal{M}}_\epsilon$. Hence,

$$\tfrac{1}{m}\sigma(\mathcal{M}) \;\leq\; \tfrac{1}{m}\sigma(\mathcal{M}_\epsilon) \;\stackrel{(2.9)}{=}\; \tfrac{1}{m}\rho(A_*^\epsilon) \;\leq\; \rho(B_\epsilon) \;\stackrel{(2.7)}{\leq}\; \rho(\mathcal{M}_\epsilon).$$

Note that $m \leq p$. Since JSR is a continuous function of the elements of the compact set $\mathcal{M}$ (see [11]), we obtain the first inequality in (2.11).

In order to prove the second inequality, assume that $\mathcal{M}$ contains an irreducible matrix A_0. Then the minimizer $x(A_0)$ of function $\phi_{A_0}(x)$ at the hyperplane $\mathcal{L} \stackrel{\mathrm{def}}{=} \{x \in R^n : \; \langle \mathbf{1}, x \rangle = 0\}$ is unique. Therefore, the corresponding restrictions of the level sets of this function are bounded. Hence, the objective function of problem (2.5) also has bounded restrictions of its level sets. Consequently, there exists at least one optimal solution $x(\mathcal{M})$ of problem (2.5) belonging to the hyperplane $\mathcal{L}$. Note that all components of $u(\mathcal{M}) \stackrel{\mathrm{def}}{=} e^{x(\mathcal{M})}$ are positive. Therefore,

$$A^T u(\mathcal{M}) \leq \sigma(\mathcal{M}) \cdot u(\mathcal{M}), \quad A \in \mathcal{M}. \tag{2.12}$$

Define now the vector norm

$$\|x\| \;=\; \sum\limits_{i=1}^n u^{(i)}(\mathcal{M}) \cdot |x^{(i)}|.$$

Then, for any $x \in R^n$ and any $A \in \mathcal{M}$ we have

$$\|Ax\| \;\leq\; \|A\,|x|\,\| \;=\; \langle u(\mathcal{M}), A|x| \rangle \;\stackrel{(2.12)}{\leq}\; \sigma(\mathcal{M}) \cdot \|x\|.$$

This means that $\rho(\mathcal{M}) \leq \sigma(\mathcal{M})$.

If $\mathcal{M}$ does not contain an irreducible matrix, we can consider the sets $\mathcal{M}_\epsilon$ with $\epsilon > 0$. By the above reasoning, we have justified that that

$$\rho(\mathcal{M}) \;\leq\; \rho(\mathcal{M}_\epsilon) \;\leq\; \sigma(\mathcal{M}_\epsilon).$$

It remains to apply Lemma 1. $\square$

Let us look now at two examples.

Example 1 *Consider $\mathcal{M} = \{A_1, \dots, A_n\}$, where the matrix $A_i = \mathbf{1} \cdot e_i^T$. Note that*

$$A_i \cdot A_j \;=\; A_j$$

for any i, $j = 1, \dots, n$. Therefore, $\rho(\mathcal{M}) = \rho(A_i) = 1$. On the other hand,

$$\begin{aligned} \sigma(\mathcal{M}) &= \inf_{u > \mathbf{0}} \max_{1 \le i \le n} \max_{1 \le j \le n} \tfrac{1}{u^{(j)}} \langle A_i e_j, u \rangle \\ &= \inf_{u > \mathbf{0}} \max_{1 \le j \le n} \tfrac{1}{u^{(j)}} \langle \mathbf{1}, u \rangle \stackrel{(2.2)}{=} \rho(\mathbf{1} \cdot \mathbf{1}^T) \;=\; n. \end{aligned}$$

Hence, the lower bound in inequality (2.11) cannot be improved. It is interesting that in this example the bound provided by $\sigma_T(\mathcal{M})$ is exact:

$$\begin{aligned} 1 &= \rho(\mathcal{M}) \stackrel{(2.11)}{\le} \sigma_T(\mathcal{M}) \\ &\stackrel{(2.8)}{=} \inf_{u > \mathbf{0}} \max_{1 \le i \le n} \max_{1 \le j \le n} \tfrac{1}{u^{(j)}} \langle A_i^T e_j, u \rangle \\ &= \inf_{u > \mathbf{0}} \max_{1 \le i \le n} \max_{1 \le j \le n} \tfrac{u^{(i)}}{u^{(j)}} \;\le\; 1. \end{aligned}$$

□

For small dimensions the JCR can be characterized in a very transparent way.

Example 2 *Let $\mathcal{M}$ be a finite set of nonnegative two-by-two matrices:*

$$\mathcal{M} \;=\; \{ A_k = (a_k, b_k), \; k = 1, \dots, m \}.$$

Then

$$\begin{aligned} \sigma(\mathcal{M}) &= \inf_{x \in R^2} \max_{1 \le k \le m} \phi_{A_k}(x) \;=\; \inf_{u > 0} \max_{1 \le k \le m} \max \left\{ \tfrac{\langle a_k, u \rangle}{u^{(1)}}, \; \tfrac{\langle b_k, u \rangle}{u^{(2)}} \right\} \\ &= \inf_{u > 0} \max \left\{ \max_{1 \le k \le m} \tfrac{\langle a_k, u \rangle}{u^{(1)}}, \; \max_{1 \le k \le m} \tfrac{\langle b_k, u \rangle}{u^{(2)}} \right\} \\ \left(\tau \stackrel{\text{def}}{=} \tfrac{u^{(2)}}{u^{(1)}} \right) &= \min_{\tau > 0} \max \left\{ \max_{1 \le k \le m} \left(a_k^{(1)} + a_k^{(2)} \cdot \tau \right), \; \max_{1 \le k \le m} \left(b_k^{(1)} \cdot \tfrac{1}{\tau} + b_k^{(2)} \right) \right\}. \end{aligned}$$

Thus, the value of $\sigma(\mathcal{M})$ is easy to find after an appropriate sorting of the coefficients. Note that this solution is a nontrivial function of the initial data. □

It is interesting that sometimes, for a richer set of variants in the switching system, it is possible to ensure that the JCR and the JSR coincide. We give a nontrivial example of such a situation in the next section.

3 Sets with independent column uncertainty

Combining the result of Lemma 2 with the upper bound (2.11), and the observation (2.7), we can see that the inclusion $A_* \in \widehat{\mathcal{M}}$ guarantees that the JSR and the JCR coincide. The simplest way to ensure this inclusion is to assume that the set $\mathcal{M}$ has independent column uncertainties:

$$\mathcal{M} \;=\; \{(a_1,\dots,a_n):\; a_i \in Q_i,\; i=1,\dots,n\}, \tag{3.1}$$

where all sets $Q_i \subset R^n$, $i=1,\dots,n$, are closed and bounded.

Theorem 2 *For a set $\mathcal{M}$ satisfying condition (3.1), we have*

$$\max_{A\in\widehat{\mathcal{M}}} \rho(A) = \rho(\mathcal{M}) = \sigma(\mathcal{M}). \tag{3.2}$$

If the solution of problem (2.5) do exist, then $\sigma(\mathcal{M}) = \rho(A_*)$ for some extreme point A_* of the set $\mathcal{M}$. Therefore,

$$\max_{A\in\mathcal{M}} \rho(A) = \rho(\mathcal{M}) = \sigma(\mathcal{M}) = \rho(A_*). \tag{3.3}$$

Proof:
Let us fix some $\epsilon > 0$. Under assumption (3.1), the matrix A_*^ϵ in (2.9) belongs to the set $\widehat{\mathcal{M}}_\epsilon$. Therefore,

$$\rho(\mathcal{M}_\epsilon) \;\stackrel{(2.11)}{\leq}\; \sigma(\mathcal{M}_\epsilon) \;\stackrel{(2.9)}{=}\; \rho(A_*^\epsilon) \;\stackrel{(2.7)}{\leq}\; \rho(\mathcal{M}_\epsilon).$$

It remains to use the continuity arguments.

Thus, we have proved that

$$\max_{A\in\widehat{\mathcal{M}}_\epsilon} \rho(A) = \rho(\mathcal{M}_\epsilon) = \sigma(\mathcal{M}_\epsilon) = \rho(A_*^\epsilon).$$

Let us show that it is always possible to choose A_*^ϵ as an extreme point of the set $\mathcal{M}_\epsilon$. For that, let us look more carefully at the structure of optimality condition for problem (2.10) written for the set $\mathcal{M}_\epsilon$. Denote $\hat{Q}_i^\epsilon \stackrel{\mathrm{def}}{=} \mathrm{Conv}\,(Q_i^\epsilon)$, $i=1,\dots,n$. Then the KKT-conditions look as follows:

$$\begin{aligned}
\mathbf{0} \;&=\; \sum_{i=1}^n \lambda_*^{(i)} g_i^*, \quad \lambda_* \in \Delta_n,\\
g_i^* \in \partial\psi_i(x_*) \;&=\; \left\{ g = \nabla\left(e^{-x^{(i)}} \langle a, e^x\rangle \right):\; a \in B_i^*(x_*) \right\},\\
B_i^*(x_*) \;&=\; \mathrm{Arg}\max_{a\in\hat{Q}_i^\epsilon} \langle a, e^{x^*}\rangle, \quad i=1,\dots,n.
\end{aligned} \tag{3.4}$$

It is clear that each $B_i^*(x_*)$ contains at least one extreme point from Q_i^ϵ.

Let us show first, that $\lambda_* > \mathbf{0}$. Note that g_i^* can be represented as

$$g_i^* \;=\; e^{-x_*^{(i)}} D(a_i^*) \cdot e^{x_*} - e^{-x_*^{(i)}} \cdot \langle a_i^*, e^{x_*}\rangle \cdot e_i$$

with certain $a_i^* \in \widehat{Q}_i^\epsilon$. Therefore,

$$\mathbf{0} \;=\; D(\sum_{i=1}^n \lambda_*^{(i)} e^{-x_*^{(i)}} a_i^*) \cdot e^{x_*} - \left(\sum_{i=1}^n \lambda_*^{(i)} e^{-x_*^{(i)}} \cdot e_i \cdot (a_i^*)^T \right) e^{x_*}.$$

Denote $A_* = (a_1^*, \ldots, a_n^*) \in \widehat{\mathcal{M}}_\epsilon$, and $\hat{A}_* = D(e^{x_*}) A_* D(e^{-x_*})$. Then, the above first-order optimality conditions can be written as follows

$$\begin{aligned} \mathbf{0} &= \left(D(e^{x_*}) A_* D(e^{-x_*}) D(\lambda_*) - D(\lambda_*) D(e^{-x_*}) A_*^T D(e^{x_*}) \right) \cdot \mathbf{1} \\ &= (\hat{A}_* D(\lambda_*) - D(\lambda_*) \hat{A}_*^T) \cdot \mathbf{1}. \end{aligned}$$

Since $A_* > \mathbf{0}$, all elements of the matrix $\hat{A}_*$ are also positive. Therefore, the above equation is possible only with $\lambda_* > \mathbf{0}$.

Let us fix an arbitrary number k, $1 \leq k \leq n$. Without loss of generality, assume that $k = 1$. Let us fix all $g_i^* \in \partial \psi_i(x_*)$, $i = 2, \ldots, n$. It is convenient to represent them in a matrix form:

$$(g_2^*, \ldots, g_n^*) \stackrel{\mathrm{def}}{=} G = - \begin{pmatrix} b^T \\ C \end{pmatrix},$$

where $G \in R^{n \times (n-1)}$ and $C \in R^{(n-1)\times(n-1)}$. Note that each $g_i \in \partial \psi_i(x_*)$ can be represented as follows:

$$g_i = \sum_{j=1}^n a^{(j)} e^{x_*^{(j)} - x_*^{(i)}} (e_j - e_i), \quad a \in B_i^*(x^*) \subset \operatorname{int} R_+^n.$$

Note that $\langle \mathbf{1}, g_i \rangle = 0$, and $g_i^{(j)} > 0$ for $j \neq i$ with $g_i^{(i)} < 0$. Therefore, matrix C is strictly diagonal dominant and it has negative off-diagonal entries. Hence, the matrix C^{-1} do exists and it has all elements positive. Therefore, the (over-determined) linear system

$$g_1 = G\lambda$$

has positive solution for any $g_1 \in \partial \psi_1(x^*)$ (the first equation of this system is a linear consequence of the others). This implies that in the optimality conditions (3.4) we can choose g_1^* using an arbitrary $a_1 \in B_1^*(x^*)$. In particular, we can choose it as an extreme point of Q_1^ϵ.

This reasoning can be repeated for the remaining indices $k = 2, \ldots, n$. Then, we end up with an extreme point $A_*^\epsilon \in \mathcal{M}_\epsilon$. Note that the latter set is obtained from $\mathcal{M}$ by adding the same ϵ to all entries of the matrices. Hence, when ϵ goes to zero, the shape of Q_1^ϵ remains unchanged. Therefore, in view of continuity of function $\rho(A)$, any limit point of extreme points A_*^ϵ is an extreme point of the set $\mathcal{M}$. □

Besides its direct applications, the Theorem 2 has an interesting algebraic consequence.

Corollary 1 *Consider the function $\rho(a_1, \ldots, a_n)$ with $a_i \in R_+^n$, $i = 1, \ldots, n$. Then this function is quasi-convex in each a_i, when all other columns are fixed.*

Proof:
Let us fix arbitrary positive vectors $x, y, a_2, \dots, a_n$. Define the uncertainty set

$$\mathcal{M} \;\; = \;\; \{A(\alpha) = (\alpha x + (1-\alpha)y, a_2, \dots, a_n), \; \alpha \in [0,1]\}.$$

Then, by Theorem 2, $\rho(\mathcal{M}) = \max\{\rho(A(0)), \rho(A(1))\}$. If the above vectors are non-negative, we can justify the result by continuity arguments. $\square$

4 Applications

Let us now consider two applications of our results.

4.1 Leontief model with uncertain data

In the input-output static Leontief model, we have n industries with production levels

$$p^{(i)} \;\; \geq \;\; 0, \quad i = 1, \dots, n.$$

In order to produce one unit of the product of industry i, we need to spend $A^{(i,j)}$ units of the product of industry j. Thus, the structure of production dependencies is given by nonnegative *consumption matrix* $A \in M_n^+$. Further, given the demand vector $d \in R_+^n$, we can find the necessary level of production as a solution to the system of linear equations:

$$p \;\; = \;\; Ap + d. \tag{4.1}$$

The economics is called *productive* if the equation (4.1) has a non-negative solution for any demand vector d. The standard condition for that is of course

$$\rho(A) \;\; < \;\; 1. \tag{4.2}$$

Clearly, the smaller is $\rho(A)$, the more flexible is the economics, and the smaller is the production level that is necessary for satisfying the current demand pattern. However, in practice the spectral radius $\rho(A)$ is difficult to estimate. Indeed, from the statistics of the previous years, we can only get sets of *different* consumption matrices

$$\mathcal{M} \;\; = \;\; \{A_1, \dots, A_k\}.$$

Hence, the joint spectral radius $\rho(\mathcal{M})$ could be a good and robust measure of economic flexibility. However, in general this value is difficult to compute even for small dimensions.

Now we have an alternative way to address this problem. Indeed, for each industry i, using the data of the previous years, we can get examples of the *distribution patterns*.[3] Denoting the convex hull of these vectors by Q_i, we can form the uncertainty set by (3.1). In this case, by Theorem 2, $\rho(\mathcal{M}) = \sigma(\mathcal{M})$, and it can be efficiently computed.

[3]Alternatively, we can use statistics on the *consumption patterns*, which corresponds to the uncertain rows of matrix A.

4.2 Optimal growth of a linear system

In some applications we need to find a matrix from a certain set, which has maximal spectral radius. If the number of matrices in the set is large, or infinite, such a problem looks completely intractable. However, it appears that in the case of independent column (or row) uncertainties we can solve the problem by aplying Theorem 2. For example, the boolean problem

$$\max_{x \in \{0,1\}^n} \rho(AD(x) + BD(\mathbf{1} - x))),$$

where the coefficients of matrices A and B are nonnegative, can be solved by convex programming techniques. Note that for each column we can consider several variants. A typical application of this type is the design of a network structure with several independent variants of connections for each node. Multiplication of an input vector by such a matrix represents the output of the system. In this case, the best design corresponds to the matrix with the largest possible spectral radius.

References

[1] M.A.Berger, Y.Wang. Bounded semigroups of matrices. Linear Algebra and its Applications, **166**, 21-27 (1992).

[2] Vincent D. Blondel, Yurii Nesterov. Computationally efficient approximations of the joint spectral radius. SIAM Journal of Matrix Analysis and Applications, **27** (1), 256-272 (2005).

[3] Vincent D. Blondel, Yurii Nesterov, Jacques Theys. On the accuracy of the ellipsoid norm approximation of the joint spectral radius. Linear Algebra and its Applications, **394**, 91-107 (2005).

[4] Vincent D. Blondel, Vincent Canterini, Undecidable problems for probabilistic automata of fixed dimension, Theory of Computing systems, **36**, 231-245, (2003).

[5] Vincent D. Blondel, Raphaël Jungers, Vladimir Protasov. On the complexity of computing the capacity of codes that avoid forbidden difference patterns. IEEE Transactions on Information Theory, **52**:(11), 5122-5127, (2006).

[6] Vincent D. Blondel, John N. Tsitsiklis. The boundedness of all products of a pair of matrices is undecidable, Systems and Control Letters, **41**:2, 135-140, (2000).

[7] D. Collela, D. Heil. Characterization of scaling functions. i. continuous solutions. SIAM Journal on Matrix Analysis and Applications, **15**, 496-518, (1994).

[8] V. Crespi, G. V. Cybenko, G. Jiang. The Theory of Trackability with Applications to Sensor Networks. Technical Report TR2005-555, Dartmouth College, Computer Science, Hanover, NH, August 2005.

[9] I. Daubechies, J. C. Lagarias. Two-scale difference equations. ii. local regularity, infinite products of matrices and fractals. SIAM Journal of Mathematical Analysis, **23**, 1031–1079, 1992.

[10] R. Horn, C. Johnson. *Matrix Analysis.* Cambridge University Press, 1985.

[11] Raphaël M. Jungers. PhD thesis, 2008.

[12] Raphaël M. Jungers, Vladimir Protasov, Vincent D. Blondel, Overlap-free words and spectra of matrices. Submitted.

[13] Raphaël Jungers, Vladimir Protasov, Vincent Blondel. Efficient algorithms for deciding the type of growth of products of integer matrices. Linear Algebra and its Applications, 428:10, 2296-2311, 2008

[14] B. E. Moision, A. Orlitsky, and P. H. Siegel. On codes that avoid specified differences. IEEE Transactions on Information Theory, 47, 433–442, 2001.

[15] Yu. Nesterov. *Introductory Lectures on Convex Optimization.* Kluwer, Boston, 2004.

[16] P.A. Parrilo, A. Jadbabaie. Approximation of the joint spectral radius using sum of squares. Linear Algebra and its Applications. 428:10, 2385-2402, 2008.

[17] V. Y. Protasov. The generalized spectral radius: A geometric approach. Izvestiya Mathematika. 61:5, 995-1030, 1997.

[18] G.-C. Rota, W. G. Strang. A note on the joint spectral radius. Indag. Math., 22, pp. 379-381, 1960.

[19] John N. Tsitsiklis, Vincent D. Blondel. The Lyapunov exponent and joint spectral radius of pairs of matrices are hard – when not impossible – to compute and to approximate. Mathematics of Control, Signals, and Systems, 10, 31-40, 1997.

Recent titles

CORE Discussion Papers

2007/93. Gaetano BLOISE and Filippo L. CALCIANO. A characterization of inefficiency in stochastic overlapping generations economies.
2007/94. Pierre DEHEZ. Shapley compensation scheme.
2007/95. Helmuth CREMER, Pierre PESTIEAU and Maria RACIONERO. Unequal wages for equal utilities.
2007/96. Helmuth CREMER, Jean-Marie LOZACHMEUR and Pierre PESTIEAU. Collective annuities and redistribution.
2007/97. Mohammed BOUADDI and Jeroen V.K. ROMBOUTS. Mixed exponential power asymmetric conditional heteroskedasticity.
2008/1. Giorgia OGGIONI and Yves SMEERS. Evaluating the impact of average cost based contracts on the industrial sector in the European emission trading scheme.
2008/2. Oscar AMERIGHI and Giuseppe DE FEO. Privatization and policy competition for FDI.
2008/3. Wlodzimierz SZWARC. On cycling in the simplex method of the Transportation Problem.
2008/4. John-John D'ARGENSIO and Frédéric LAURIN. The real estate risk premium: A developed/emerging country panel data analysis.
2008/5. Giuseppe DE FEO. Efficiency gains and mergers.
2008/6. Gabriella MURATORE. Equilibria in markets with non-convexities and a solution to the missing money phenomenon in energy markets.
2008/7. Andreas EHRENMANN and Yves SMEERS. Energy only, capacity market and security of supply. A stochastic equilibrium analysis.
2008/8. Géraldine STRACK and Yves POCHET. An integrated model for warehouse and inventory planning.
2008/9. Yves SMEERS. Gas models and three difficult objectives.
2008/10. Pierre DEHEZ and Daniela TELLONE. Data games. Sharing public goods with exclusion.
2008/11. Pierre PESTIEAU and Uri POSSEN. Prodigality and myopia. Two rationales for social security.
2008/12. Tim COELLI, Mathieu LEFEBVRE and Pierre PESTIEAU. Social protection performance in the European Union: comparison and convergence.
2008/13. Loran CHOLLETE, Andréas HEINEN and Alfonso VALDESOGO. Modeling international financial returns with a multivariate regime switching copula.
2008/14. Filomena GARCIA and Cecilia VERGARI. Compatibility choice in vertically differentiated technologies.
2008/15. Juan D. MORENO-TERNERO. Interdependent preferences in the design of equal-opportunity policies.
2008/16. Ana MAULEON, Vincent VANNETELBOSCH and Wouter VERGOTE. Von Neumann-Morgenstern farsightedly stable sets in two-sided matching.
2008/17. Tanguy ISAAC. Information revelation in markets with pairwise meetings: complete information revelation in dynamic analysis.
2008/18. Juan D. MORENO-TERNERO and John E. ROEMER. Axiomatic resource allocation for heterogeneous agents.
2008/19. Carlo CAPUANO and Giuseppe DE FEO. Mixed duopoly, privatization and the shadow cost of public funds.
2008/20. Helmuth CREMER, Philippe DE DONDER, Dario MALDONADO and Pierre PESTIEAU. Forced saving, redistribution and nonlinear social security schemes.
2008/21. Philippe CHEVALIER and Jean-Christophe VAN DEN SCHRIECK. Approximating multiple class queueing models with loss models.
2008/22. Pierre PESTIEAU and Uri M. POSSEN. Interaction of defined benefit pension plans and social security.
2008/23. Marco MARINUCCI. Optimal ownership in joint ventures with contributions of asymmetric partners.

Recent titles

CORE Discussion Papers - continued

2008/24. Raouf BOUCEKKINE, Natali HRITONENKO and Yuri YATSENKO. Optimal firm behavior under environmental constraints.
2008/25. Ana MAULEON, Vincent VANNETELBOSCH and Cecilia VERGARI. Market integration in network industries.
2008/26. Leonidas C. KOUTSOUGERAS and Nicholas ZIROS. Decentralization of the core through Nash equilibrium.
2008/27. Jean J. GABSZEWICZ, Didier LAUSSEL and Ornella TAROLA. To acquire, or to compete? An entry dilemma.
2008/28. Jean-Sébastien TRANCREZ, Philippe CHEVALIER and Pierre SEMAL. Probability masses fitting in the analysis of manufacturing flow lines.
2008/29. Marie-Louise LEROUX. Endogenous differential mortality, non monitored effort and optimal non linear taxation.
2008/30. Santanu S. DEY and Laurence A. WOLSEY. Two row mixed integer cuts via lifting.
2008/31. Helmuth CREMER, Philippe DE DONDER, Dario MALDONADO and Pierre PESTIEAU. Taxing sin goods and subsidizing health care.
2008/32. Jean J. GABSZEWICZ, Didier LAUSSEL and Nathalie SONNAC. The TV news scheduling game when the newscaster's face matters.
2008/33. Didier LAUSSEL and Joana RESENDE. Does the absence of competition *in* the market foster competition *for* the market? A dynamic approach to aftermarkets.
2008/34. Vincent D. BLONDEL and Yurii NESTEROV. Polynomial-time computation of the joint spectral radius for some sets of nonnegative matrices.

Books

Y. POCHET and L. WOLSEY (eds.) (2006), *Production planning by mixed integer programming*. New York, Springer-Verlag.
P. PESTIEAU (ed.) (2006), *The welfare state in the European Union: economic and social perspectives*. Oxford, Oxford University Press.
H. TULKENS (ed.) (2006), *Public goods, environmental externalities and fiscal competition*. New York, Springer-Verlag.
V. GINSBURGH and D. THROSBY (eds.) (2006), *Handbook of the economics of art and culture*. Amsterdam, Elsevier.
J. GABSZEWICZ (ed.) (2006), La différenciation des produits. Paris, La découverte.
L. BAUWENS, W. POHLMEIER and D. VEREDAS (eds.) (2008), *High frequency financial econometrics: recent developments*. Heidelberg, Physica-Verlag.
P. VAN HENTENRYCKE and L. WOLSEY (eds.) (2007), *Integration of AI and OR techniques in constraint programming for combinatorial optimization problems*. Berlin, Springer.

CORE Lecture Series

C. GOURIÉROUX and A. MONFORT (1995), Simulation Based Econometric Methods.
A. RUBINSTEIN (1996), Lectures on Modeling Bounded Rationality.
J. RENEGAR (1999), A Mathematical View of Interior-Point Methods in Convex Optimization.
B.D. BERNHEIM and M.D. WHINSTON (1999), Anticompetitive Exclusion and Foreclosure Through Vertical Agreements.
D. BIENSTOCK (2001), Potential function methods for approximately solving linear programming problems: theory and practice.
R. AMIR (2002), Supermodularity and complementarity in economics.
R. WEISMANTEL (2006), Lectures on mixed nonlinear programming.